

The Faiss library

Year: 2024 | Citations: 360 | Authors: Matthijs Douze, Alexandr Guzhva, Chengqi Deng, Jeff Johnson, Gergely Szilvassy

Abstract

Vector databases typically manage large collections of embedding vectors. Currently, AI applications are growing rapidly, and so is the number of embeddings that need to be stored and indexed. The Faiss library is dedicated to vector similarity search, a core functionality of vector databases. Faiss is a toolkit of indexing methods and related primitives used to search, cluster, compress and transform vectors. This paper describes the trade-off space of vector search and the design principles of Faiss in terms of structure, approach to optimization and interfacing. We benchmark key features of the library and discuss a few selected applications to highlight its broad applicability.